

The Computing of Digital Ecosystems

Gerard Briscoe

Media and Communications

London School of Economics and Political Science

United Kingdom

Philippe De Wilde

Department of Computer Science

Heriot Watt University

United Kingdom

Abstract

A primary motivation for our research in *digital ecosystems* is the desire to exploit the self-organising properties of biological ecosystems. Ecosystems are thought to be robust, scalable architectures that can automatically solve complex, dynamic problems. However, the computing technologies that contribute to these properties have not been made explicit in *digital ecosystems* research. Here, we discuss how different computing technologies can contribute to providing the necessary self-organising features, including Multi-Agent Systems (MASs), Service-Oriented Architectures (SOAs), and distributed evolutionary computing (DEC). The potential for exploiting these properties in *digital ecosystems* is considered, suggesting how several key features of biological ecosystems can be exploited in Digital Ecosystems, and discussing how mimicking these features may assist in developing robust, scalable self-organising architectures. An example architecture, the Digital Ecosystem, is considered in detail. The Digital Ecosystem is then measured experimentally through simulations, considering the self-organised diversity of its evolving agent populations relative to the user request behaviour.

1 Introduction

Digital Ecosystems are distributed adaptive open socio-technical systems, with properties of self-organisation, scalability and sustainability, inspired by natural ecosystems [4, 8, 5], and are emerging as a novel approach to the catalysis of sustainable regional development driven by Small and Medium sized Enterprises (SMEs). Digital Ecosystems aim to help local economic actors become active players in globalisation, valorising their local culture and vocations, and enabling them to interact and create value networks at the global level [14, 43].

We have previously considered the biological inspiration for the technical component of Digital Ecosystems [10, 9], and we will now consider the relevant computing technologies. Based on our understanding of biological ecosystems in the context of Digital Ecosystems [6], we will now introduce fields from the domain of computer science relevant to the creation of Digital Ecosystems. As we are interested in the digital counterparts for the behaviour and constructs of biological ecosystems, instead of simulating or emulating such behaviour or constructs, we will consider what parallels can be drawn. We will start by considering MASs to explore the references to *agents* and *migration* [10, 9]; followed by evolutionary computing and SOAs for the references to *evolution* and *self-organisation* [10, 9].

The value of creating parallels between biological and computer systems varies substantially depending on the behaviours or constructs being compared, and sometimes cannot

be done so convincingly. For example, both have mechanisms to ensure data integrity. In computer systems, that integrity is absolute, data replication which introduces even the most minor change is considered to have failed, and is supported by mechanisms such as the Message-Digest algorithm 5 [39]. While in biological systems, the genetic code is transcribed with a remarkable degree of fidelity; there is, approximately, only one unforced error per one hundred bases copied [32]. There are also elaborate proof-reading and correction systems, which in evolutionary terms are highly conserved [32]. In this example establishing a parallel is infeasible, despite the relative similarity in function, because the operational control mechanisms in biological and computing systems are radically different, as are the aims and purposes. This is a reminder that considerable finesse is required when determining parallels, or when using existing ones.

2 Multi-Agent Systems

A *software agent* is a piece of software that acts, for a user in a relationship of *agency*, autonomously in an environment to meet its designed objectives. A MAS is a system composed of several *software agents*, collectively capable of reaching goals that are difficult to achieve by an individual agent or monolithic system. Conceptually, there is a strong parallel between the software agents of a MAS and the agent-based models of a biological ecosystem [15], despite the lack of evolution and migration in a MAS. There is an even stronger parallel to a variant of MASs, called *mobile agent systems*, in which the mobility also mirrors the migration in biological ecosystems [38].

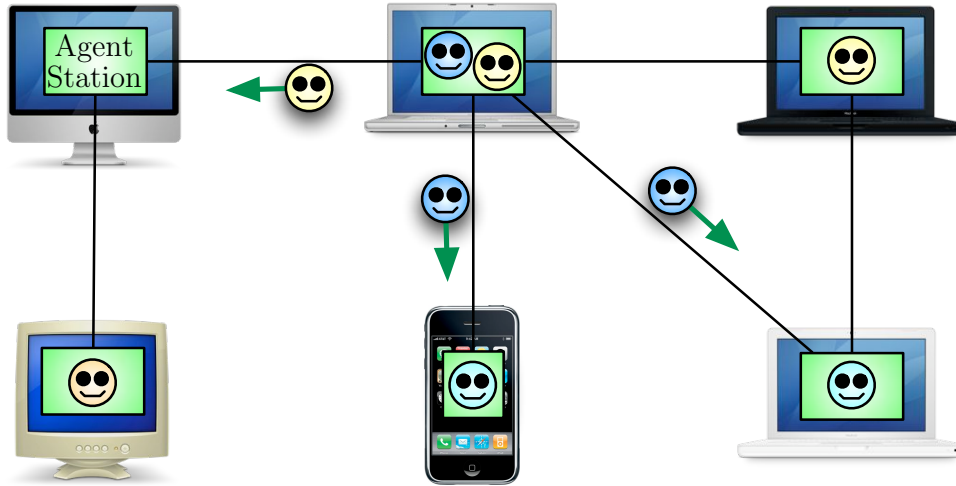


Figure 1: *Mobile Agent System: Visualisation that shows mobile agents as programmes that can migrate from one host to another in a network of heterogeneous computer systems and perform a task specified by its owner.*

The term *mobile agent* contains two separate and distinct concepts: mobility and agency [40]. Hence, mobile agents are software agents capable of movement within a network [38]. The mobile agent paradigm proposes to treat a network as multiple agent-*friendly* environments and the agents as programmatic entities that move from location to location, performing tasks for users. So, on each host they visit mobile agents need software which is responsible for their execution, providing a safe execution environment [38].

Generally, there are three types of design for mobile agent systems [38]: (1) using a specialised operating system, (2) as operating system services or extensions, or (3)

as application software. The third approach builds mobile agent systems as specialised application software that runs on top of an operating system, to provide for the mobile agent functionality, with such software being called an *agent station* [31]. In this last approach, each agent station hides the vendor-specific aspects of its host platform, and offers standardised services to visiting agents. Services include access to local resources and applications; for example, web servers or *web services*, the local exchange of information between agents via message passing, basic security services, and the creation of new agents [31]. Also, the third approach is the most platform-agnostic, and is visualised in Figure 1.

3 Service-Oriented Architectures

To evolve high-level software components in Digital Ecosystems, we propose taking advantage of the *native* method of software advancement, human developers, and the use of *evolutionary computing* for *combinatorial optimisation* of the available software services. This involves treating developer-produced software services as the functional building blocks, as the base unit in a genetic-algorithms-based process. Our approach to evolving high-level software applications requires a modular reusable paradigm to software development, such as SOAs. SOAs are the current state-of-the-art approach, being the current iteration of interface/component-based design from the 1990s, which was itself an iteration of event-oriented design from the 1980s, and before then modular programming from the 1970s [22]. Service-oriented computing promotes assembling application components into a loosely coupled network of services, to create flexible, dynamic business processes and agile applications that span organisations and computing platforms [37]. This is achieved through a SOA, an architectural style that guides all aspects of creating and using business processes throughout their life-cycle, packaged as services. This includes defining and provisioning infrastructure that allows different applications to exchange data and participate in business processes, loosely coupled from the operating systems and programming languages underlying the applications. Hence, a SOA represents a model in which functionality is decomposed into distinct units (services), which can be distributed over a network, and can be combined and reused to create business applications [37].

A SOA depends upon service-orientation as its fundamental design principle. In a SOA environment, independent services can be accessed without knowledge of their underlying platform implementation. Services reflect a *service-oriented* approach to programming that is based on composing applications by discovering and invoking network-available services to accomplish some task. This approach is independent of specific programming languages or operating systems, because the services communicate with each other by passing data from one service to another, or by co-ordinating an activity between two or more services [37]. So, the concepts of SOAs are often seen as built upon, and the development of, the concepts of modular programming and distributed computing [22].

SOAs allow for an information system architecture that enables the creation of applications that are built by combining loosely coupled and interoperable services. They typically implement functionality most people would recognise as a service, such as filling out an online application for an account, or viewing an online bank statement [22]. Services are intrinsically unassociated units of functionality, without calls to each other embedded in them. Instead of services embedding calls to each other in their source code, protocols are defined which describe how services can talk to each other, in a process known as orchestration, to meet new or existing business system requirements. This is allowing an increasing number of third-party software companies to offer software services, such that SOA systems will come

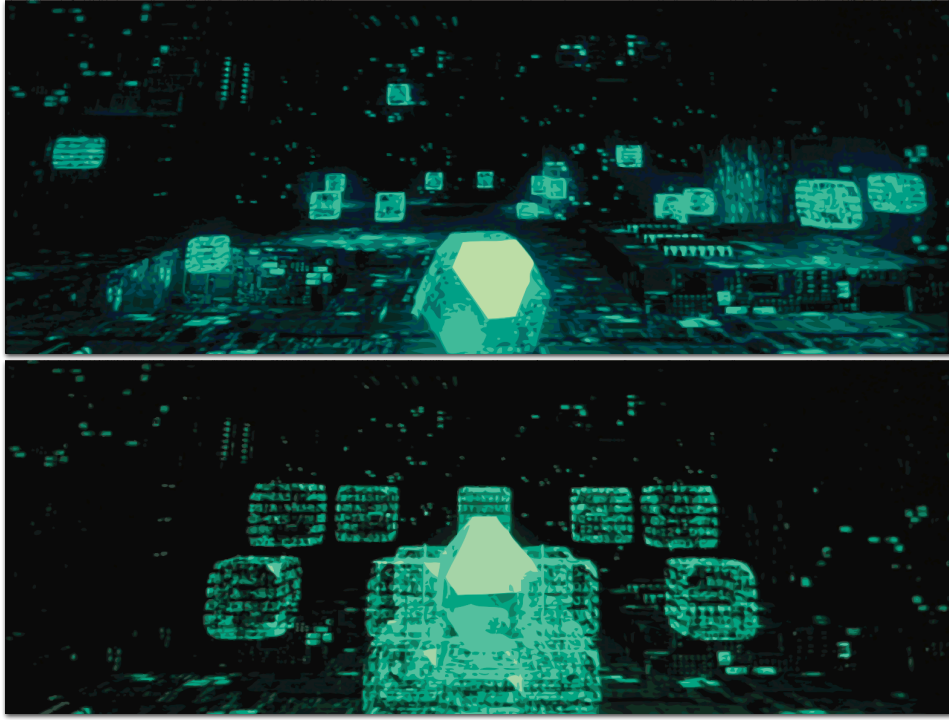


Figure 2: *Service-Oriented Architectures: Abstract visualisations, with the loosely joined services as cuboids, and the service orchestration as a polyhedron.*

to consist of such third-party services combined with others created in-house, which has the potential to spread costs over many users and uses, and promote standardisation both in and across industries. For example, the travel industry now has a well-defined, and documented, set of both services and data, sufficient to allow any competent software engineer to create travel agency software using entirely off-the-shelf software services [13]. Other industries, such as the finance industry, are also making significant progress in this direction [48].

The vision of SOAs assembling application components from a loosely coupled network of services, that can create dynamic business processes and agile applications that span organisations and computing platforms, is visualised in Figure 2. It will be made possible by creating compound solutions that use internal organisational software assets, including enterprise information and legacy systems, and combining these solutions with external components residing in remote networks. The great promise of SOAs is that the *marginal cost* of creating the n -th application is virtually zero, as all the software required already exists to satisfy the requirements of other applications. Only their *combination* and *orchestration* are required to produce a new application [33]. The *key* is that the interactions between the *chunks* are not specified within the *chunks* themselves. Instead, the interaction of services (all of whom are hosted by unassociated peers) is specified by users in an ad-hoc way, with the intent driven by newly emergent business requirements [27].

The pinnacle of SOA interoperability, is the exposing of services on the internet as *web services*. A web service is a specific type of service that is identified by a Uniform Resource Identifier (URI), whose service description and transport utilise open Internet standards. Interactions between web services typically occur as Simple Object Access Protocol (SOAP) calls carrying eXtensible Markup Language (XML) data content. The interface descriptions of web services are expressed using the Web Services Definition Language (WSDL) [36]. The Universal Description Discovery and Integration (UDDI) standard defines a protocol for

directory services that contain web service descriptions. UDDI enables web service clients to locate candidate services and discover their details. Service clients and service providers utilise these standards to perform the basic operations of SOAs [36]. Service aggregators can then use the Business Process Execution Language (BPEL) to create new web services by defining corresponding compositions of the interfaces and internal processes of existing services [36].

SOA services inter-operate based on a formal definition (or contract, e.g. WSDL) that is independent of the underlying platform and programming language. Service descriptions are used to advertise the service capabilities, interface, behaviour, and quality [36]. The publication of such information about available services provides the necessary means for discovery, selection, binding, and composition of services [36]. The (expected) behaviour of a service during its execution is described by its behavioural description (for example, as a workflow process). Also, included is a quality of service (QoS) description, which publishes important functional and non-functional service quality attributes, such as service metering and cost, performance metrics (response time, for instance), security attributes, integrity (transactional), reliability, scalability, and availability [36]. Service clients (end-user organisations that use some service) and service aggregators (organisations that consolidate multiple services into a new, single service offering) utilise *service descriptions* to achieve their objectives [36]. One of the most important and continuing developments in SOAs is Semantic Web Services (SWS), which make use of *semantic descriptions* for service discovery, so that a client can discover the services semantically [11].

There are multiple standards available and still being developed for SOAs, most notably of recent being REpresentational State Transfer (REST) [42]. The software industry now widely implements a thin SOAP/WSDL/UDDI veneer atop existing applications or components that implement the web services paradigm, but the choice of technologies will change with time. Therefore, the fundamentals of SOAs and its services are best defined generically, because SOAs are technology agnostic and need not be tied to a specific technology [37]. Within the current and future scope of the fundamentals of SOAs, there is clearly potential to *evolve* complex high-level software applications from the modular services of SOAs, instead of the instruction level evolution currently prevalent in genetic programming [21].

4 Distributed Evolutionary Computing

Having previously suggested evolutionary computing [10, 9], and the possibility of it occurring within a distributed environment, not unlike those found in mobile agent systems, leads us to consider a specialised form known as DEC. The fact that evolutionary computing manipulates a population of independent solutions actually makes it well suited for parallel and distributed computation architectures [12]. The motivation for using parallel or distributed evolutionary algorithms is twofold: first, improving the speed of evolutionary processes by conducting concurrent evaluations of individuals in a population; second, improving the problem-solving process by overcoming difficulties that face traditional evolutionary algorithms, such as maintaining diversity to avoid premature convergence [44]. There are several variants of distributed evolutionary computing, leading some to propose a taxonomy for their classification, with there being two main forms [12, 44]:

- multiple-population/coarse-grained migration/island
- single-population/fine-grained diffusion/neighbourhood

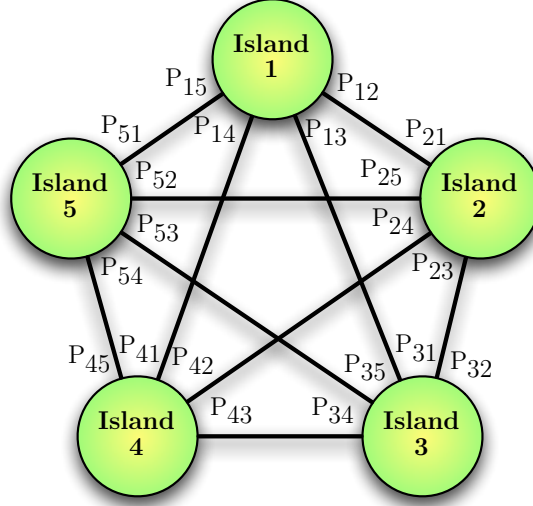


Figure 3: *Island-Model of Distributed Evolutionary Computing: There are different probabilities of going from island 1 to island 2, as there is of going from island 2 to island 1.*

In the coarse-grained *island* models [28, 12], evolution occurs in multiple parallel sub-populations (islands), each running a local evolutionary algorithm, evolving independently with occasional *migrations* of highly fit individuals among sub-populations. The core parameters for the evolutionary algorithm of the island-models are as follows [28]:

- number of the sub-populations: 2, 3, 4, more
- sub-population homogeneity
 - size, crossover rate, mutation rate
- topology of connectivity: ring, star, fully-connected
- static or dynamic connectivity
- migration mechanisms:
 - isolated/synchronous/asynchronous
 - how often migrations occur
 - which individuals migrate

Fine-grained *diffusion* models [44] assign one individual per processor. A local neighbourhood topology is assumed, and individuals are allowed to mate only within their neighbourhood, called a *deme*. The demes overlap by an amount that depends on their shape and size, and in this way create an implicit migration mechanism. Each processor runs an identical evolutionary algorithm which selects parents from the local neighbourhood, produces an offspring, and decides whether to replace the current individual with an offspring. However, even with the advent of multi-processor computers, and more recently multi-core processors, which provide the ability to execute multiple threads simultaneously, this approach would still prove impractical in supporting the number of agents necessary to create a Digital Ecosystem. Therefore, we shall further consider the *island* models.

An example island-model [28, 12] is visualised in Figure 3, in which there are different probabilities of going from island 1 to island 2, as there is of going from island 2 to island

1. This allows maximum flexibility for the migration process, and mirrors the naturally inspired quality that although two populations have the same physical separation, it may be easier to migrate in one direction than the other, i.e. fish migration is easier downstream than upstream. The migration of the *island* models is like the notion of migration in nature, being similar to the metapopulation models of theoretical ecology [26]. This model has also been used successfully in the determination of investment strategies in the commercial sector, in a product known as the Galapagos toolkit [45]. However, all the *islands* in this approach work on exactly the same problem, which makes it less analogous to biological ecosystems in which different locations can be environmentally different [3]. We will take advantage of this property later when defining the Ecosystem-Oriented Architecture of Digital Ecosystems.

5 The Digital Ecosystem

We will now define the architectural principles of Digital Ecosystems. We will use our understanding of theoretical biology [6, 4], mimicking the processes and structures of life, evolution, and ecology of biological ecosystems. We will achieve this by combining elements from mobile agents systems, distributed evolutionary computing, and Service-Oriented Architectures, to create a hybrid architecture which is the *digital counterpart of biological ecosystems*. Combining these technologies, based on the biological inspiration, the technical component of Digital Ecosystems provide a two-level optimisation scheme inspired by natural ecosystems, in which a decentralised peer-to-peer network forms an underlying tier of distributed agents. These agents then feed a second optimisation level based on an evolutionary algorithm that operates locally on single habitats (peers), aiming to find solutions that satisfy locally relevant constraints. The local search is sped up through this twofold process, providing better local optima as the distributed optimisation provides prior sampling of the search space by making use of computations already performed in other peers with similar constraints. So, the Digital Ecosystem supports the automatic combining of numerous agents (which represent services), by their interaction in evolving populations to meet user requests for applications, in a scalable architecture of distributed interconnected habitats. The sharing of agents between habitats ensures the system is scalable, while maintaining a high evolutionary specialisation for each user. The network of interconnected habitats is equivalent to the *abiotic* environment of biological ecosystems [3]; combined with the agents, the populations, the agent migration for distributed evolutionary computing, and the environmental selection pressures provided by the user base, then the union of the habitats creates the Digital Ecosystem, which is summarised in Figure 4. The continuous and varying user requests for applications provide a dynamic evolutionary pressure on the applications (agent aggregations), which have to evolve to better fulfil those user requests, and without which there would be no driving force to the evolutionary self-organisation of the Digital Ecosystem.

If we consider an example user base for the Digital Ecosystem, the use of SOAs in its definition means that business-to-business (B2B) interaction scenarios [22] lend themselves to being a potential user base for Digital Ecosystems. So, we can consider a *business ecosystem* of Small and Medium sized Enterprise (SME) networks [34], as a specific class of examples for B2B interaction scenarios; and in which the SME users are requesting and providing software services, represented as agents in the Digital Ecosystem, to fulfil the needs of their business processes, creating a Digital Business Ecosystem as shown in Figure 5. SOAs promise to provide potentially huge numbers of services that programmers can combine, via the standardised interfaces, to create increasingly more sophisticated and distributed

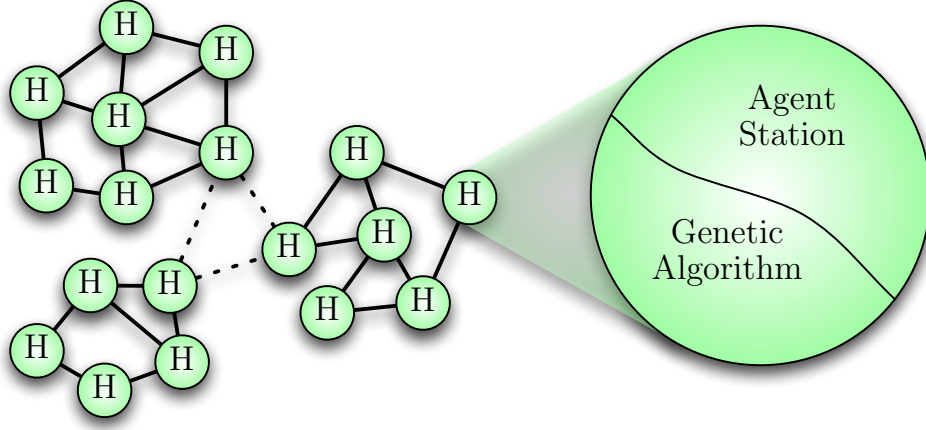


Figure 4: *Digital Ecosystem: Optimisation architecture in which agents (representing services) travel along the P2P connections; in every node (habitat) local optimisation is performed through an evolutionary algorithm, where the search space is determined by the agents present at the node.*

applications [36]. The Digital Ecosystem extends this concept with the automatic combining of available and applicable services, represented by agents, in a scalable architecture, to meet user requests for applications. These agents will recombine and evolve over time, constantly seeking to improve their effectiveness for the user base. From the SME users' point of view the Digital Ecosystem provides a network infrastructure where connected enterprises can advertise and search for services (real-world or software only), putting a particular emphasis on the composability of loosely coupled services and their optimisation to local and regional, needs and conditions. To support these SME users the Digital Ecosystem is satisfying the companies' business requirements by finding the most suitable services or combination of services (applications) available in the network. An application (composition of services) is defined as an agent aggregation (collection) in the habitat network that can move from one peer (company) to another, being hosted only in those where it is most useful in satisfying the SME users' business needs.

The agents consist of an *executable component* and an *ontological description*. So, the Digital Ecosystem can be considered a MAS which uses *distributed evolutionary computing* [12, 44] to combine suitable agents in order to meet user requests for applications.

The landscape, in energy-centric biological ecosystems, defines the connectivity between habitats [3]. Connectivity of nodes in the digital world is generally not defined by geography or spatial proximity, but by information or semantic proximity. For example, connectivity in a peer-to-peer network is based primarily on bandwidth and information content, and not geography. The island-models of distributed evolutionary computing use an information-centric model for the connectivity of nodes (*islands*) [28]. However, because it is generally defined for one-time use (to evolve a solution to one problem and then stop) it usually has a fixed connectivity between the nodes, and therefore a fixed topology [12]. So, supporting evolution in the Digital Ecosystem, with a multi-objective *selection pressure* (fitness landscape with many peaks), requires a re-configurable network topology, such that habitat connectivity can be dynamically adapted based on the observed migration paths of the agents between the users within the habitat network. Based on the island-models of distributed evolutionary computing [28], each connection between the habitats is bi-directional and there is a probability associated with moving in either direction across the

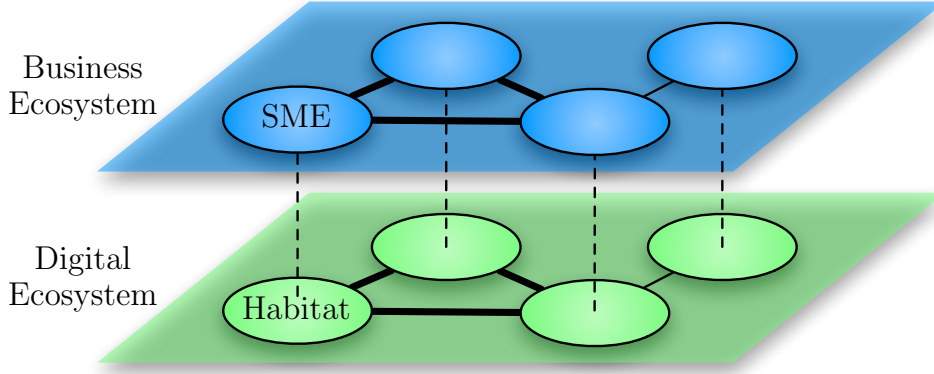


Figure 5: *Digital Business Ecosystem: Business ecosystem, network of SMEs, using the Digital Ecosystem. The habitat clustering will therefore be parallel to the business sector communities.*

connection, with the connection probabilities affecting the rate of migration of the agents. However, additionally, the connection probabilities will be updated by the success or failure of agent migration using the concept of Hebbian learning: the habitats which do not successfully exchange agents will become less strongly connected, and the habitats which do successfully exchange agents will achieve stronger connections. This leads to a topology that adapts over time, resulting in a network that supports and resembles the connectivity of the user base. If we consider a *business ecosystem*, network of SMEs, as an example user base; such business networks are typically small-world networks [47]. They have many strongly connected clusters (communities), called *sub-networks* (quasi-complete graphs), with a few connections between these clusters (communities). Graphs with this topology have a very high clustering coefficient and small characteristic path lengths [46]. So, the Digital Ecosystem will take on a topology similar to that of the user base, as shown in Figure 5.

The novelty of our approach comes from the evolving populations being created in response to *similar* requests. So whereas in the island-models of distributed evolutionary computing there are multiple evolving populations in response to one request [28], here there are multiple evolving populations in response to *similar* requests. In our Digital Ecosystems different requests are evaluated on separate *islands* (populations), and so adaptation is accelerated by the sharing of solutions between evolving populations (islands), because they are working to solve similar requests (problems).

The users will formulate queries to the Digital Ecosystem by creating a request as a *semantic description*, like those being used and developed in SOAs, specifying an application they desire and submitting it to their local peer (habitat). This description defines a metric for evaluating the *fitness* of a composition of agents, as a distance function between the *semantic description* of the request and the agents' *ontological descriptions*. A population is then instantiated in the user's habitat in response to the user's request, seeded from the agents available at their habitat. This allows the evolutionary optimisation to be accelerated in the following three ways: first, the habitat network provides a subset of the agents available globally, which is localised to the specific user it represents; second, making use of applications (agent aggregations) previously evolved in response to the user's earlier requests; and third, taking advantage of relevant applications evolved elsewhere in response to similar requests by other users. The population then proceeds to evolve the optimal application (agent aggregation) that fulfils the user request, and as the agents are the base unit for

evolution, it searches the available agent combination space. For an evolved agent aggregation (application) that is executed by the user, it then migrates to other peers (habitats) becoming hosted where it is useful, to combine with other agents in other populations to assist in responding to other user requests for applications.

6 Simulation and Results

We have previously compared the performance of Digital Ecosystems to traditional SOA based systems, and found that it outperforms at greater scales [6], and so here we focus on the self-organising behaviour which makes this possible. Self-organisation is perhaps one of the most desirable features in the systems that we design, and a primary motivation for our research in Digital Ecosystems is the desire to exploit the self-organising properties of biological ecosystems [25], which are thought to be robust, scalable architectures that can automatically solve complex, dynamic problems. Over time a biological ecosystem becomes increasingly self-organised through the process of *ecological succession* [3], driven by the evolutionary self-organisation of the populations within the ecosystem. Analogously, a Digital Ecosystem’s increasing self-organisation comes from the agent populations within being evolved to meet the dynamic *selection pressures* created by the requests from the user base. The self-organisation of biological ecosystems is often defined in terms of the *complexity*, *stability*, and *diversity* [19], which we will also apply to our Digital Ecosystems, here considering the *diversity*.

Self-organisation has been around since the late 1940s [1], but has escaped general formalisation despite many attempts [35, 20]. There have instead been many notions and definitions of self-organisation, useful within their different contexts [17]. They have come from cybernetics [1, 2, 18], thermodynamics [35], mathematics [24], information theory [41], synergetics [16], and other domains [23]. The term *self-organising* is widely used, but there is no generally accepted meaning, as the abundance of definitions would suggest. Therefore, the philosophy of self-organisation is complicated, because organisation has different meanings to different people. So, we would argue that any definition of self-organisation is context dependent, in the same way that a choice of statistical measure is dependent on the data being analysed.

While we could measure the self-organised *diversity* of individual evolving agent populations, or even take a random sampling, it will be more informative to consider their *collective* self-organised *diversity*. Also, given that the Digital Ecosystem is required to support a range of user behaviour, we can consider the *collective* self-organised *diversity* of the evolving agent populations relative to the *global* user request behaviour. So, when varying a behavioural property of the user requests according to some distribution, we would expect the corresponding property of the evolving agent populations to follow the same distribution. We are not intending to prescribe the expected user behaviour of the Digital Ecosystem, but investigate whether the Digital Ecosystem can adapt to a range of user behaviour. So, we will consider Uniform, Gaussian (Normal) and Power Law distributions for the parameters of the user request behaviour. The Uniform distribution will provide a control, while the Normal (Gaussian) distribution will provide a reasonable assumption for the behaviour of a large group of users, and the Power Law distribution will provide a relatively extreme variation in user behaviour. So, we simulated the Digital Ecosystem, varying aspects of the user behaviour according to different distributions, and measuring the related aspects of the evolving agent populations. This consisted of a mechanism to vary the user request properties of *length* and *modularity* (number of attributes per atomic service), according to

Uniform, Gaussian (normal) and Power Law distributions, and a mechanism to measure the corresponding application (agent aggregation) properties of *size* and *number of attributes per agent*. For statistical significance each scenario (experiment) will be averaged from ten thousand simulation runs. We expect it will be obvious whether the *observed* behaviour of the Digital Ecosystem matches the *expected* behaviour from the user base. Nevertheless, we will also implement a chi-squared (χ^2) test to determine if the observed behaviour (distribution) of the agent aggregation properties matches the expected behaviour (distribution) from the user request properties.

6.1 User Request Length

We started by varying the *user request length* according to the available distributions, expecting the size of the corresponding applications (agent aggregations) to be distributed according to the length of the user requests, i.e. the longer the user request, the larger the agent aggregation needed to fulfil it.

We first applied the Uniform distribution as a control, and graphed the results in Figure 6. The *observed* frequencies of the application (agent aggregation) size mostly matched the *expected* frequencies, which was *confirmed* by a χ^2 test; with a *null hypothesis* of *no significant difference* and *sixteen degrees of freedom*, the χ^2 value was 2.588, below the critical 0.95 χ^2 value of 7.962.

We then applied the Gaussian distribution as a reasonable assumption for the behaviour of a large group of users, and graphed the results in Figure 7. The *observed* frequencies of the application (agent aggregation) size matched the *expected* frequencies with only minor variations, which was confirmed by a χ^2 test; with a *null hypothesis* of *no significant difference* and *sixteen degrees of freedom*, the χ^2 value was 2.102, below the critical 0.95 χ^2 value of 7.962.

Finally, we applied the Power Law distribution to represent a relatively extreme variation in user behaviour, and graphed the results in Figure 8. The *observed* frequencies of the application (agent aggregation) size matched the *expected* frequencies with some variation, which was confirmed by a χ^2 test; with a *null hypothesis* of *no significant difference* and *sixteen degrees of freedom*, the χ^2 value was 5.048, below the critical 0.95 χ^2 value of 7.962.

There were a couple of minor discrepancies, similar to all the experiments. First, there were a small number of *individual* agents at the thousandth time step, caused by the typical user behaviour of continuously creating new services (agents). Second, while the chi-squared tests confirmed that there was no significant difference between the *observed* and *expected* frequencies of the application (agent aggregation) size, there was still a *bias* to larger applications (solutions). Evident visually in the graphs of the experiments, and evident numerically in the chi-squared test of the Power Law distribution experiment as it favoured shorter agent-sequences. The cause of this *bias* was most likely some aspect of *bloat* not fully controlled.

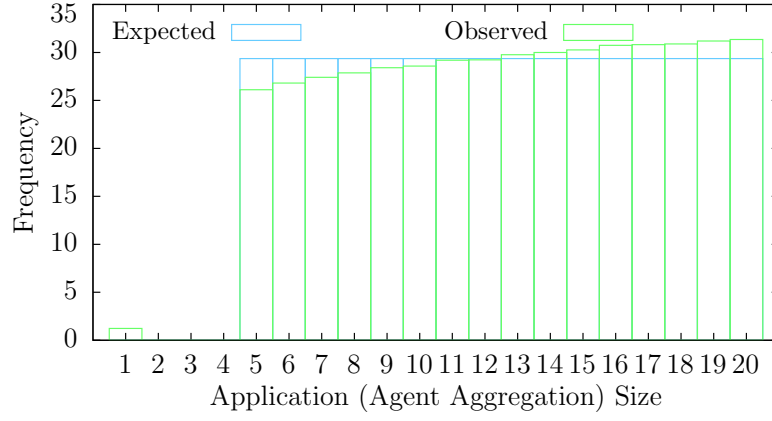


Figure 6: *Graph of Uniformly Distributed Agent-Sequence Length Frequencies: The observed frequencies of the application (agent aggregation) size mostly matched the expected frequencies.*

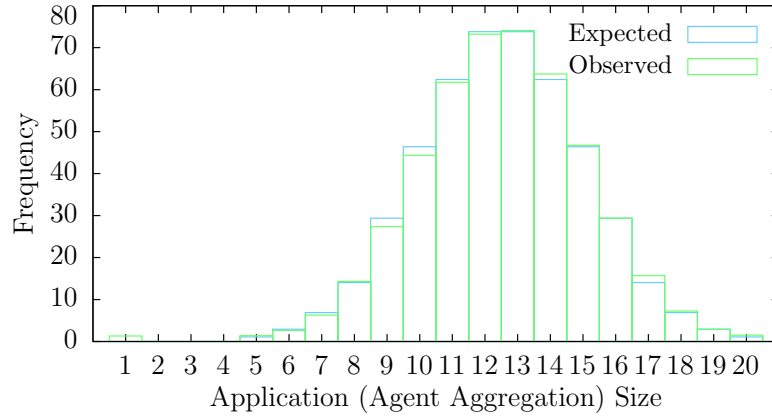


Figure 7: *Graph of Gaussian Distributed Agent-Sequence Length Frequencies: The observed frequencies of the application (agent aggregation) size matched the expected frequencies with only minor variations.*

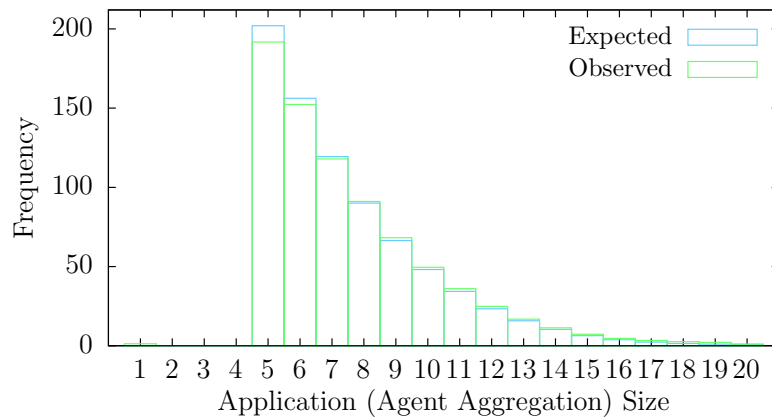


Figure 8: *Graph of Power Law Distributed Agent-Sequence Length Frequencies: The observed frequencies of the application (agent aggregation) size matched the expected frequencies with some variation.*

6.2 User Request Modularity

Next, we varied the *user request modularity* (number of attributes per atomic service) according to the available distributions, expecting the *sophistication* of the agents to be distributed according to the modularity of the user requests, i.e. the more complicated (in terms of modular non-reducible tasks) the user request, the more sophisticated (in terms of the number of attributes) the agents needed to fulfil it.

We first applied the Uniform distribution as a control, and graphed the results in Figure 9. The *observed* frequencies for the number of agent attributes mostly matched the *expected* frequencies, which was confirmed by a χ^2 test; with a *null hypothesis* of *no significant difference* and *ten degrees of freedom*, the χ^2 value was 1.049, below the critical 0.95 χ^2 value of 3.940.

We then applied the Gaussian distribution as a reasonable assumption for the behaviour of a large group of users, and graphed the results in Figure 10. The *observed* frequencies for the number of agent attributes again followed the *expected* frequencies, but there was variation which led to a failed χ^2 test; with a *null hypothesis* of *no significant difference* and *ten degrees of freedom*, the χ^2 value was 50.623, not below the critical 0.95 χ^2 value of 3.940.

Finally, we applied the Power Law distribution to represent a relatively extreme variation in user behaviour, and graphed the results in Figure 11. The *observed* frequencies for the number of agent attributes also followed the *expected* frequencies, but there was variation which led to a failed χ^2 test; with a *null hypothesis* of *no significant difference* and *ten degrees of freedom*, the χ^2 value was 61.876, not below the critical 0.95 χ^2 value of 3.940.

In all the experiments the *observed* frequencies of the number of agent attributes followed the *expected* frequencies, with some variation in two of the experiments. This could only be confirmed statistically, by a χ^2 test, for the Uniform distribution experiment. In the Gaussian and Power distribution experiments the χ^2 tests failed by considerable margins, most likely because the evolving agent populations were still self-organising to match the user behaviour, shown by the *observed* frequencies *approaching* the *expected* frequencies, but not yet sufficiently to meet χ^2 tests, because by the thousandth *time step* (user request event) each user had placed an average of only ten requests.

Collectively, the experimental results confirm that the self-organised *diversity* of the evolving agent populations is relative to the *selection pressures* of the user base, which was confirmed statistically for most of the experiments. So, we have determined an effective understanding and quantification for the self-organised *diversity* of the evolving agent populations of our Digital Ecosystem. While the minor experimental failures, in which the Digital Ecosystem responded more slowly than in the other experiments, have shown that there is potential to optimise the Digital Ecosystem, because the evolutionary self-organisation of an ecosystem is a slow process [3], even the accelerated form present in our Digital Ecosystem.

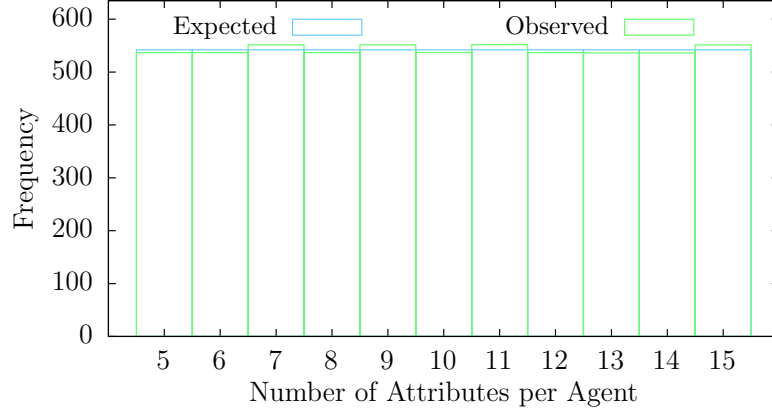


Figure 9: *Graph of Uniformly Distributed Agent Attribute Frequencies: The observed frequencies for the number of agent attributes mostly matched the expected frequencies.*

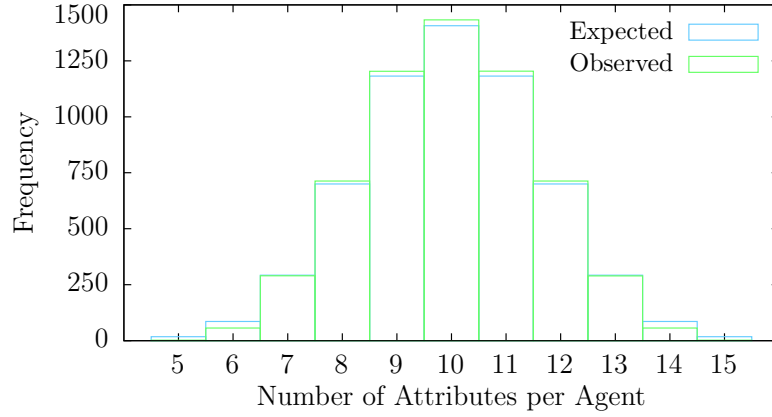


Figure 10: *Graph of Gaussian Distributed Agent Attribute Frequencies: The observed frequencies for the number of agent attributes again followed the expected frequencies, but there was variation.*

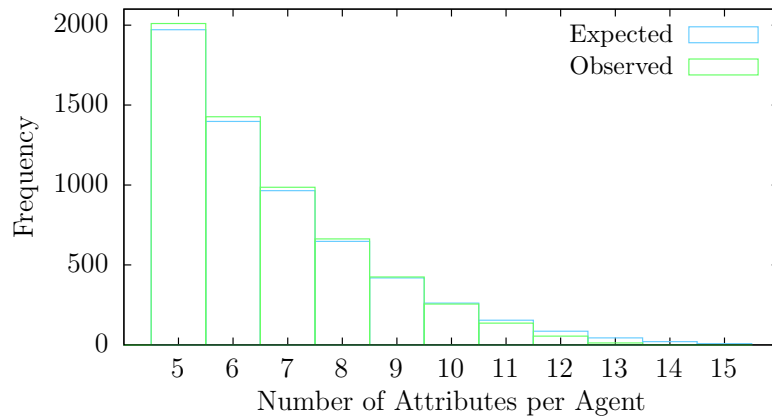


Figure 11: *Graph of Power Law Distributed Agent Attribute Frequencies: The observed frequencies for the number of agent attributes also followed the expected frequencies, but there was variation.*

7 Conclusions

We have confirmed the fundamentals for a new class of system, Digital Ecosystems, created through combining understanding from theoretical ecology, evolutionary theory, MASs, distributed evolutionary computing, and SOAs. Digital Ecosystems, where the word *ecosystem* is more than just a metaphor, being the digital counterpart of biological ecosystems, and therefore having their desirable properties, such as scalability and self-organisation. It is a complex system that shows emergent behaviour, being more than the sum of its constituent parts.

With the core infrastructure of Digital Ecosystems defined, some intelligence-based augmentations have been considered for its acceleration and optimisation [7, 4]. In which the agents were endowed with memory, intelligence and the ability to share information, with the aim of optimising the ecological dynamics, through the emergent optimisation of the agent migration patterns. While not yet as sophisticated as swarm intelligence, we would argue it is a step in that direction, and we would like to investigate further in the future.

The ever-increasing challenge of software complexity in creating progressively more sophisticated and distributed applications, makes the design and maintenance of efficient and flexible systems a growing challenge [30], for which current software development techniques have hit a *complexity wall* [29]. In response we have created Digital Ecosystems, the digital counterparts of biological ecosystems, possessing their properties of self-organisation, scalability and sustainability [25]; Ecosystem-Oriented Architectures that overcome the challenge by automating the search for new algorithms in a scalable architecture, through the evolution of software services in a distributed network.

Acknowledgments

We would like to thank for helpful discussions Paolo Dini of the London School of Economics and Political Science. This work was supported by the EU-funded Open Philosophies for Associative Autopoietic Digital Ecosystems (OPAALS) Network of Excellence (NoE), Contract No. FP6/IST-034824.

References

- [1] W Ashby. Principles of the self-organizing dynamic system. *Journal of General Psychology*, 37:125–128, 1947.
- [2] S. Beer. *Decision and Control: The Meaning of Operational Research and Management Cybernetics*. Wiley, 1966.
- [3] M Begon, J Harper, and C Townsend. *Ecology: Individuals, Populations and Communities*. Blackwell Publishing, 1996.
- [4] G Briscoe. *Digital Ecosystems*. PhD thesis, Imperial College London, 2009.
- [5] G Briscoe. Complex adaptive digital ecosystems. In *ACM Management of Emergent Digital Ecosystems Conference*, 2010.
- [6] G Briscoe and P De Wilde. Digital Ecosystems: Evolving service-oriented architectures. In *IEEE Bio Inspired Models of Network, Information and Computing Systems Conference*, 2006.

- [7] G Briscoe and P De Wilde. Digital Ecosystems: Optimisation by a distributed intelligence. In *IEEE Digital Ecosystems and Technologies Conference*, 2008. (Awarded Best Track Paper).
- [8] G Briscoe and P De Wilde. Computing of applied digital ecosystems. In *ACM Management of Emergent Digital Ecosystems Conference*, 2009.
- [9] G Briscoe and S Sadedin. Natural science paradigms. In *Digital Business Ecosystems*, pages 48–55. European Commission, 2007.
- [10] G Briscoe, S Sadedin, and G Paperin. Biology of applied digital ecosystems. In *IEEE Digital Ecosystems and Technologies Conference*, pages 458–463, 2007.
- [11] L. Cabral, J. Domingue, E. Motta, T. Payne, and F. Hakimpour. Approaches to semantic web services: an overview and comparisons. In *The Semantic Web: Research and Applications*, pages 225–239. Springer, 2004.
- [12] E. Cantu-Paz. A survey of parallel genetic algorithms. *Réseaux et systèmes répartis, Calculateurs Parallèles*, 10:141–171, 1998.
- [13] Jorge Cardoso and Amit Sheth. Introduction to semantic web services and web process composition. In Jorge Cardoso and Amit Sheth, editors, *Semantic Web Services and Web Process Composition*, pages 1–13. Springer, 2004.
- [14] P. Dini, G. Lombardo, R. Mansell, A.R. Razavi, S. Moschoyiannis, P. Krause, A. Nicolai, and L Rivera León. Beyond interoperability to digital ecosystems: regional innovation and socio-economic development led by SMEs. *International Journal of Technological Learning, Innovation and Development*, 1:410–426, 2008.
- [15] D Green, N Klomp, G Rimmington, and S Sadedin. *Complexity in Landscape Ecology*. Springer, 2006.
- [16] H. Haken. *Synergetics: an introduction: nonequilibrium phase transitions and self-organization in physics, chemistry, and biology*. Springer, 1977.
- [17] F. Heylighen. The science of self-organization and adaptivity. *The Encyclopedia of Life Support Systems*, pages 253–280, 2002.
- [18] F. Heylighen and C. Joslyn. Cybernetics and second order cybernetics. *Encyclopedia of Physical Science & Technology*, 4:155–170, 2001.
- [19] A. King and S. Pimm. Complexity, diversity, and stability: A reconciliation of theoretical and empirical results. *American Naturalist*, 122:229–239, 1983.
- [20] T. Kohonen. *Self-organization and associative memory*. Springer, 1989.
- [21] J Koza. Overview of genetic programming. In *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, pages 73–78. MIT, 1992.
- [22] D. Krafzig, K. Banke, and D. Slama. *Enterprise SOA: Service-Oriented Architecture Best Practices*. Prentice Hall, 2004.
- [23] J.M. Lehn. Perspectives in supramolecular chemistry—from molecular recognition towards molecular information processing and self-organization. *Angewandte Chemie International Edition in English*, 29:1304–1319, 1990.
- [24] G. Lendaris. On the definition of self-organizing systems. In Wolfgang Banzhaf, editor, *European Conference on Advances in Artificial Life*, pages 324–325. IEEE Press, 1964.
- [25] S Levin. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems*, 1:431–436, 1998.

- [26] R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15:237–240, 1969.
- [27] F. Leymann, D. Roller, and M. Schmidt. Web services and business process management. *IBM Systems Journal*, 41:198–211, 2002.
- [28] S. Lin, W. Punch III, and E. Goodman. Coarse-grain parallel genetic algorithms: categorization and new approach. In *Symposium on Parallel and Distributed Processing*, pages 28–37. IEEE Press, 1994.
- [29] K. Lyytinen and Y. Yoo. The next wave of nomadic computing: A research agenda for information systems research. *Sprouts: Working Papers on Information Systems*, 1:1–20, 2001.
- [30] J. Markoff. Faster chips are leaving programmers in their dust. Technical report, New York Times, 2007.
- [31] F. McCabe and K. Clark. April-agent process interaction language. In Michael Wooldridge and Nick Jennings, editors, *Intelligent Agents: Workshop on Agent Theories, Architectures, and Languages*, pages 324–340. Springer, 1994.
- [32] S. McCulloch, R. Kokoska, O. Chilkova, C. Welch, E. Johansson, P. Burgers, and T. Kunkel. Enzymatic switching for efficient and accurate translesion DNA replication. *Nucleic Acids Research*, 32:4665–4675, 2004.
- [33] Gurpreet Modi. Service oriented architecture & web 2.0. Technical report, Guru Tegh Bahadur Institute of Technology, 2007.
- [34] J. Moore. *The Death of Competition: Leadership and Strategy in the Age of Business Ecosystems*. Harvard Business School Press, 1996.
- [35] G. Nicolis and I. Prigogine. *Self-organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. Wiley, 1977.
- [36] M. Papazoglou and D. Georgakopoulos. Service-oriented computing. *Communications of the ACM*, 46:25–28, 2003.
- [37] M. Papazoglou. Service-oriented computing: concepts, characteristics and directions. In Tiziana Catarci, Massimo Mecella, John Mylopoulos, and Maria Orlowska, editors, *International Conference on Web Information Systems Engineering*, pages 3–12. IEEE Press, 2003.
- [38] V. Pham and A. Karmouch. Mobile software agents: an overview. *IEEE Communications Magazine*, 36:26–37, 1998.
- [39] R. Rivest. The MD5 message-digest algorithm. Technical report, MIT, 1992.
- [40] K. Roethermel and F. Hohl. Mobile agents. In A. Kent and J. Williams, editors, *Encyclopedia for Computer Science and Technology*, volume 40, pages 155–176. CRC Press, 1999.
- [41] C. Shalizi. *Causal Architecture, Complexity and Self-Organization in Time Series and Cellular Automata*. PhD thesis, University of Wisconsin-Madison, 2001.
- [42] M. Singh and M. Huhns. *Service-Oriented Computing: Semantics, Processes, Agents*. Wiley, 2005.
- [43] J. Stanley and G. Briscoe. The ABC of digital business ecosystems. *Communications Law - Journal of Computer, Media and Telecommunications Law*, 15(1), 2010.

- [44] J. Stender. *Parallel Genetic Algorithms: Theory and Applications*. IOS Press, 1993.
- [45] M Ward. Life offers lessons for business, 2004.
- [46] D Watts and S Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442, 1998.
- [47] D White and M. Houseman. The navigability of strong ties: Small worlds, tie strength, and network topology. *Complexity*, 8:72–81, 2002.
- [48] O. Zimmermann, S. Milinski, M. Craes, and F. Oellermann. Second generation web services-oriented architecture in production in the finance industry. In John Vlissides and Douglas Schmidt, editors, *Conference on Object-Oriented Programming, Systems, Languages, and Applications*, pages 283–289. ACM Press, 2004.